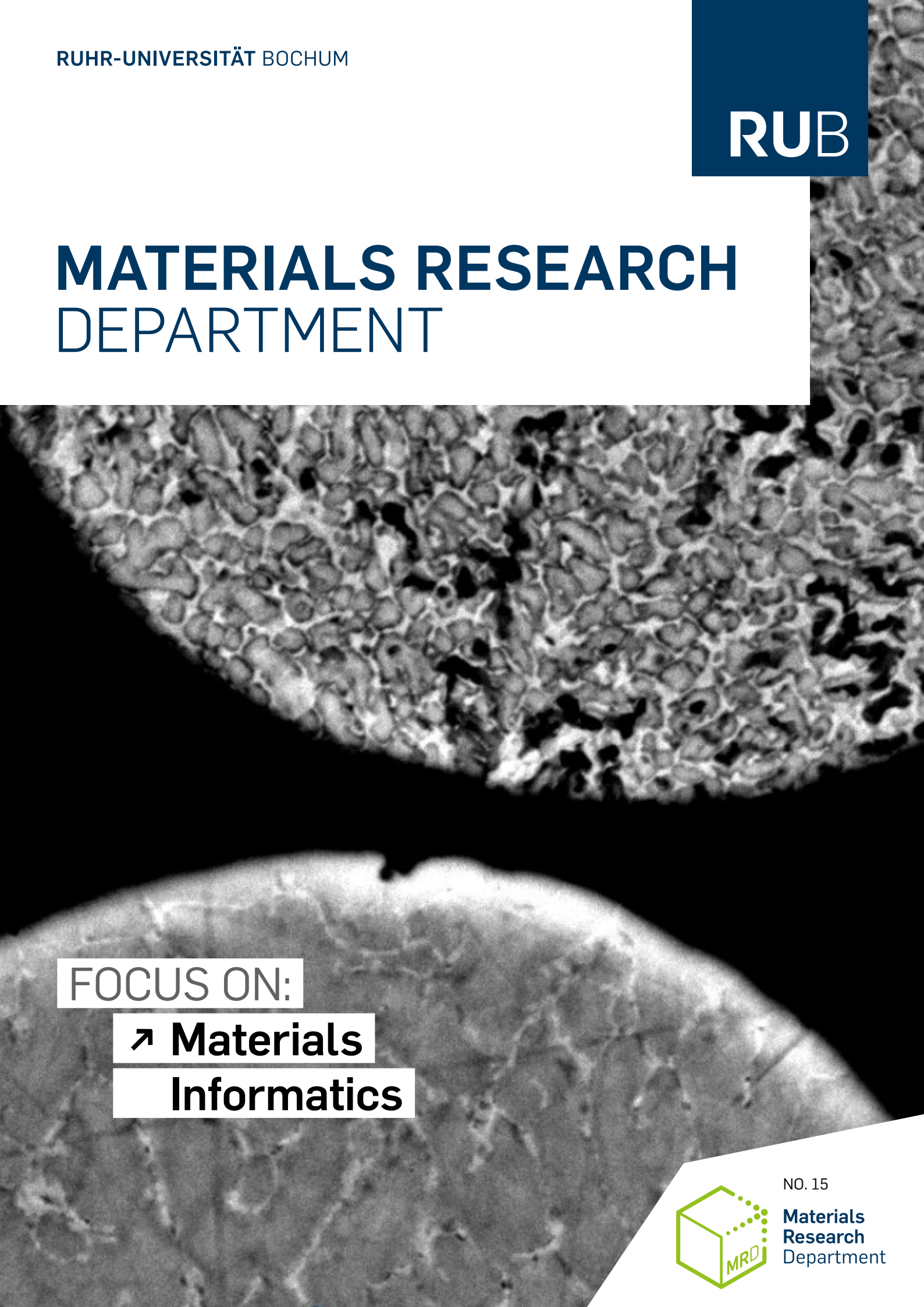
RUHR-UNIVERSITÄT BOCHUM

RUB

MATERIALS RESEARCH DEPARTMENT



FOCUS ON:

↗ Materials Informatics

NO. 15

Materials Research Department

MRD

MATERIALS RESEARCH DEPARTMENT

FOCUS TOPIC: MATERIALS INFORMATICS

Materials informatics is changing materials science completely. Data science and machine learning accelerates analysis, simulation, design and discovery of materials. Our 15th MRD newsletter highlights some of MRD's research in this quickly evolving field.

Like all previous newsletters, this 15th issue is accessible through the MRD website. We are looking forward to receiving feedback on our newsletter and we welcome suggestions for the next issues.

Enjoy reading,
R. Drautz and A. Devi
MRD Speakers
P. Aleithe and F. Scholz
MRD Science Managers

CONTENT

A Knowledge Graph for Superalloys ... 03

Microstructure Informatics ... 04

Using Convolutional Neural Networks for Fast and Reliable Microstructural Analysis ... 06

Atomistics Meets Statistics ... 08

Three-dimensional Microstructure Reconstruction from Surface EBSD Maps ... 10

Atomic Cluster Expansion for Large-scale Atomistic Simulations with Quantum Accuracy ... 12

CONTACT

Materials Research Department
Universitätsstr. 150
D-44801 Bochum

@ mrd@rub.de
↗ www.mrd.rub.de

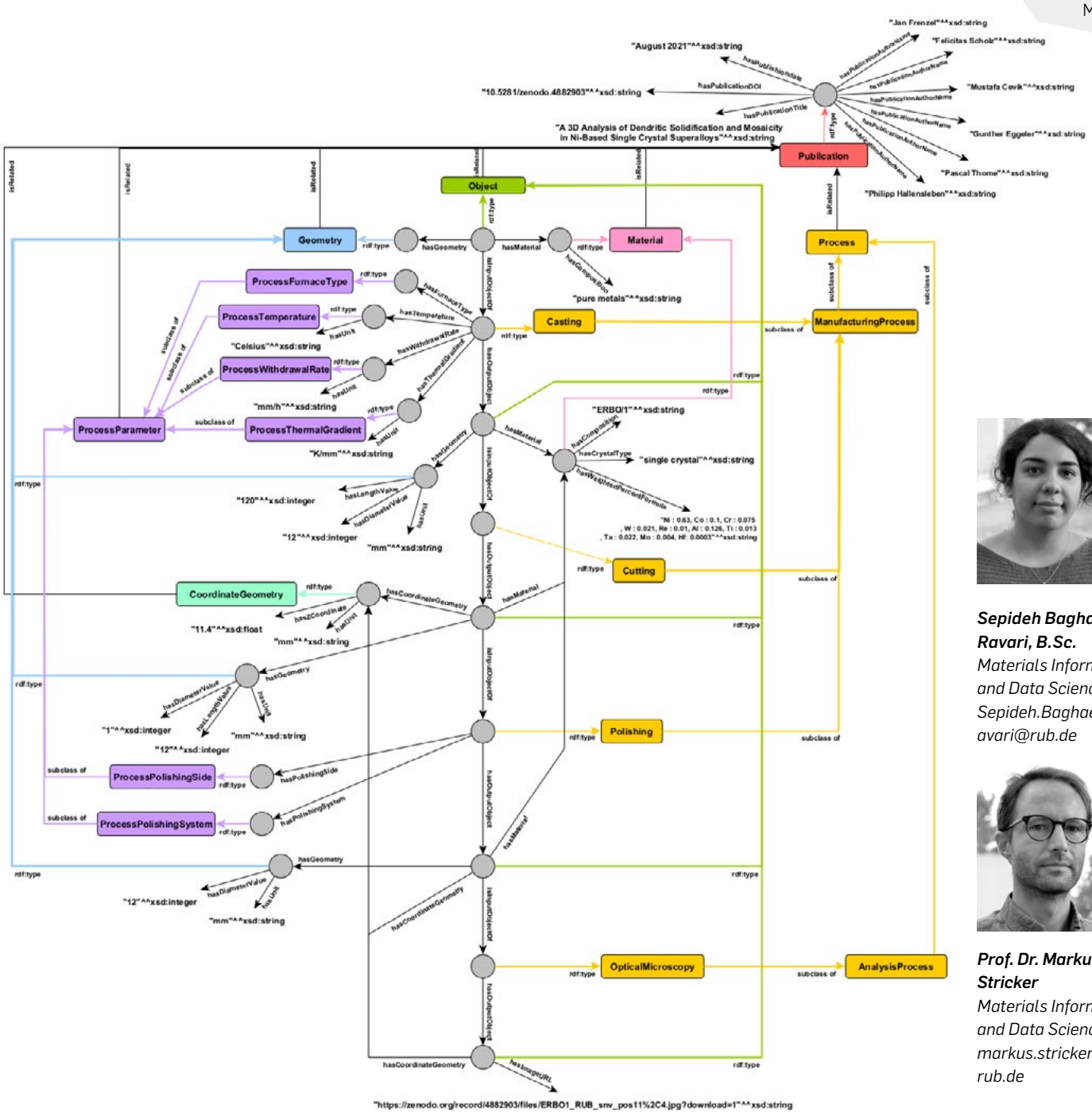
A KNOWLEDGE GRAPH FOR SUPERALLOYS

One Step in the Digitalization Process of Knowledge

The main mode of communicating new scientific knowledge was and still is the academic peer-reviewed paper. Data is presented in various forms: micrographs, functional dependencies (x/y plots), photographs, etc. Knowledge is communicated by providing an introductory context in written form with references to other papers, a presentation of the used equipment, materials, and methods (reproducibility), and the resulting measurements including their description (results), again in text form, typically followed by a discussion that extracts principles, relationships and provides generalizations. Finally, the most important findings and their consequences are concluded and possible limitations are stated. This is one of the best ways (we know of) to communicate knowledge to humans. The knowledge is, however, not machine-readable in general. "Machine-readable" refers to the possibility of (semi-)automated access to the data and contained knowledge through algorithms querying a data source.

The Collaborative Research Centre (CRC) Transregio 103 "From atoms to turbine blades – a scientific basis for a new generation of single crystal superalloys" [1] is currently in its last funding phase. Results from more than a decade of research are already published in more than 300 papers with more to come. This body of literature constitutes the scientific output of the CRC, accessible to humans. Links between the papers exist through citations, but the actual knowledge from these papers is not linked and not machine-readable and, therefore, less accessible to further scrutineering. To create a knowledge graph of the accumulated knowledge, we present one step in the digitalization process of the knowledge using already published experimental work.

A knowledge graph consists of the data itself and a data model (ontology) which defines how individual data are connected. This knowledge graph can then be queried with computers for analytics (How many measurements exist for a given alloy? Are there connections between measurements that



Sepideh Baghaee Ravari, B.Sc.
Materials Informatics and Data Science
Sepideh.BaghaeeRavari@rub.de

Prof. Dr. Markus Stricker
Materials Informatics and Data Science
markus.stricker@rub.de

have not been exploited?) but also serves as an entry point for further analysis to infer hidden latent knowledge through various Materials Informatics methods. As a first step towards the creation of a full knowledge graph of all the papers published in the CRC, we devised a protocol to formalize the knowledge contained in Scholz et al. [2] through a paper-specific ontology. The figure shows shows the developed ontology. It includes bibliographic information, chemical composition, all described processes for sample preparation, and finally the analysis process which led to the published results and conclusions. The result of this protocol is a digital representation of the data and, most importantly, how this data was obtained (metadata) by Scholz et al. [2]. But for a knowledge graph to be useful for Materials Informatics approaches, it has to contain the knowledge of many more papers.

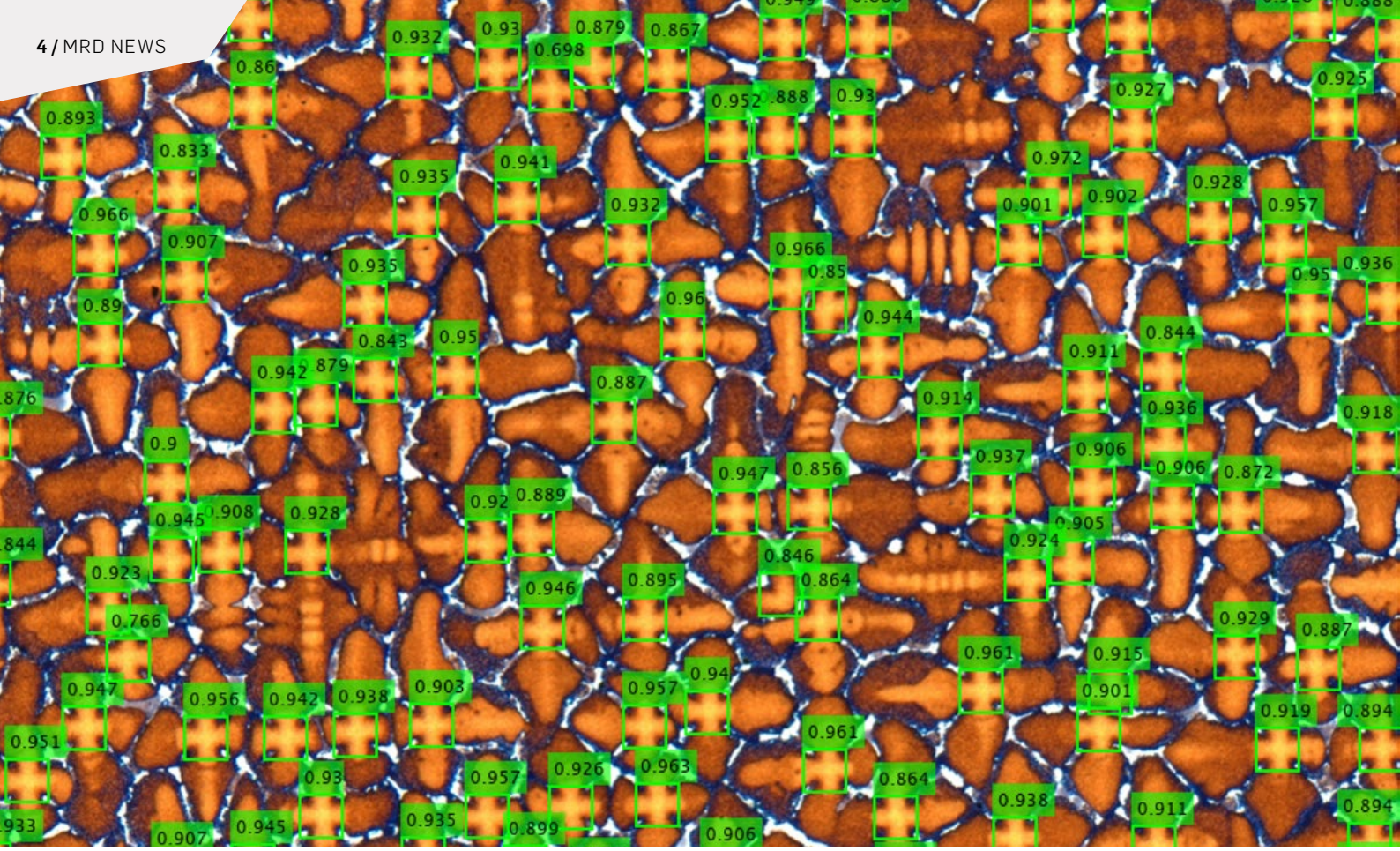
Our current protocol requires a substantial amount of manual work to read the paper, structure the information, and translate it into an ontology. This manual process is very instructive but does not scale to the over 300 articles published in the context of the CRC. Two solutions are required in the future for knowledge representation. One for already published work, as demonstrated here, and one for the future publication of findings. The former involves the development and deployment of tools from the field of natural language processing and text-parsing tools to automate the extraction of information from text-based data [3]. This allows to automatic processing of knowledge in text form which can then be included in an existing knowledge graph. The latter requires fundamental changes in how data is (experimentally) recorded and annotated through digital event logging, electronic lab notebooks, as well as a research data management system [3] and is a work in progress in many communities in Materials Science and Engineering.

REFERENCES:

[1] https://www.sfb-transregio103.de/.
[2] Scholz, F., et al. Materials (2021), 14(17), 4904.
[3] Hawizy, L., et al., Journal of Cheminformatics (2011), 3, 17.

WEB :

Link to online version of the ontology with an automatic layout
↗ **http://owlgred.lumii.lv/online_visualization/lw2w**



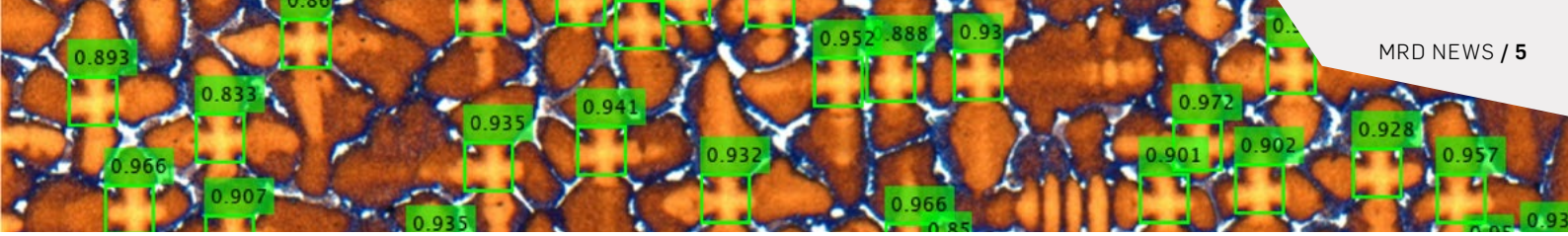
MICROSTRUCTURE INFORMATICS

Automatic Characterization of Material Morphologies with Advanced Computer Vision Techniques

Quantitative characterization methods of material microstructures represent important tools in materials science. A solid understanding of microstructural features is required to develop suitable models which help to understand the relations between manufacturing, morphology and performance of engineering materials. In recent years, the capability and reliability of advanced computer vision techniques has drastically improved [1]. Artificial neural networks, as well as machine and deep learning approaches have proven their potential in different fields of science, e.g. in applications for autonomous driving cars [2]. Similar methods can be adapted for quantitative image analysis in material scientific projects where they can be used to automatically characterize different types of microstructures with an accuracy that outperforms classical algorithmic computer vision approaches [3].

One interesting use case for automatic image analysis based on deep learning techniques is the 3D characterization of dendritic microstructures in single crystal nickel based super alloys. Directional solidification is a key processing technique for today's superalloys. The evolution of the cast microstructure is accompanied by the competitive growth of dendrites, leading to the formation of crystal mosaicity. Crystal mosaicity is characterized by small misorientations between neighboring dendrites. During solidification, interactions between growing dendrites (e.g. extinctions of primary or branching of tertiary dendrite arms) determine the morphology of the cast microstructure and can thereby influence the high temperature creep and fatigue resistance of the material.

To gain a better understanding of the geometrical details of the interactions and how they affect the as-cast microstructure, a three-dimensional



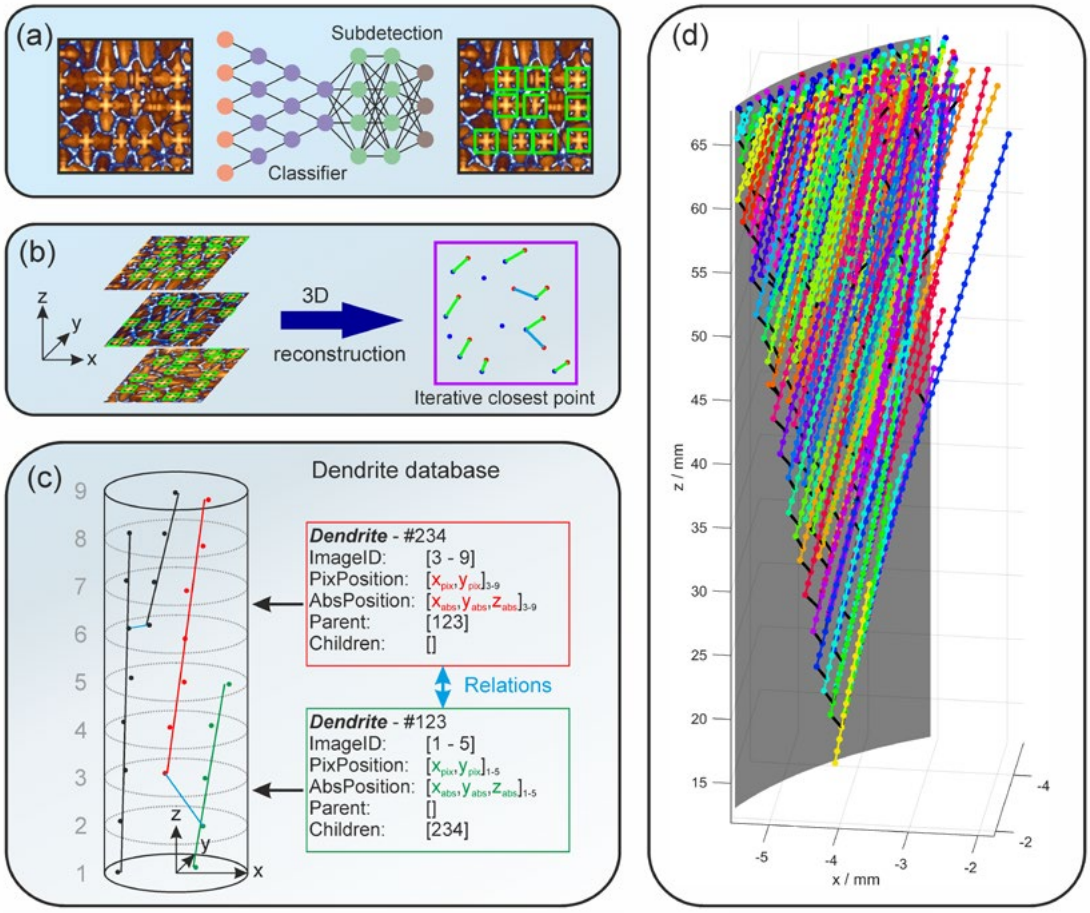


Figure 1: Architecture of a deep-learning assisted automatic 3D dendrite characterization method.

characterization process is required, where individual dendrites are traced over a certain distance of solidification [4]. For this purpose, we developed an automatic characterization routine with four characteristic features: (1) A tomographic serial sectioning procedure, where equidistant cross-sections are first cut out of the cylindrical single crystalline specimen and then characterized by optical microscopy. (2) A deep learning object detection architecture consisting of a classifier and a subdetection network, which locates the coordinates of the primary dendrite arms, Fig. 1a. (3) An image registration routine based on an iterative closest points algorithm to reconstruct the 3D growth paths of dendrites in space, Fig. 1b. And (4), a 3D dendrite specific data base capturing all geometric properties and interactions between the dendrites, Fig. 1c. In this way we effectively created a digital twin of our dendritic sample which can now be used to extract morphological information in an automated fashion. This allows us to investigate interactions between multiple dendrites as well as interactions between dendrites and the mold wall during solidification, Fig. 1d. Furthermore, overall statistical properties of the complete dendrite population in terms of the evolution of dendrite spacings, the frequency of extinctions and branching events and the effect of misaligned dendrites on their surroundings, can be extracted. This approach allows for an efficient automatic exploration of microstructural properties and sets us up for future utilization in terms of big data mining.

Alexander Richter, B.Sc.hons
Chair Materials Science and Engineering
Alexander.Richter-b6w@rub.de

Dr.-Ing. Pascal Thome
Chair Materials Science and Engineering
pascal.thome@rub.de

REFERENCES:

[1] LeCun, Y., et al., Nature (2015), 521(7553), 436-444.
[2] Devi, S., et al., Wireless Personal Communications (2020), 114(3), 2121-2133.
[3] Stuckner, J., et al., NPJ Computational Materials (2020), 8(1), 200.
[4] Scholz, F., et al. Materials (2021), 14(17), 4904.

USING **CONVOLUTIONAL NEURAL NETWORKS** FOR FAST AND RELIABLE MICRO-STRUCTURAL ANALYSIS

Exemplified by Textures of Solidification Structures in Powder Samples

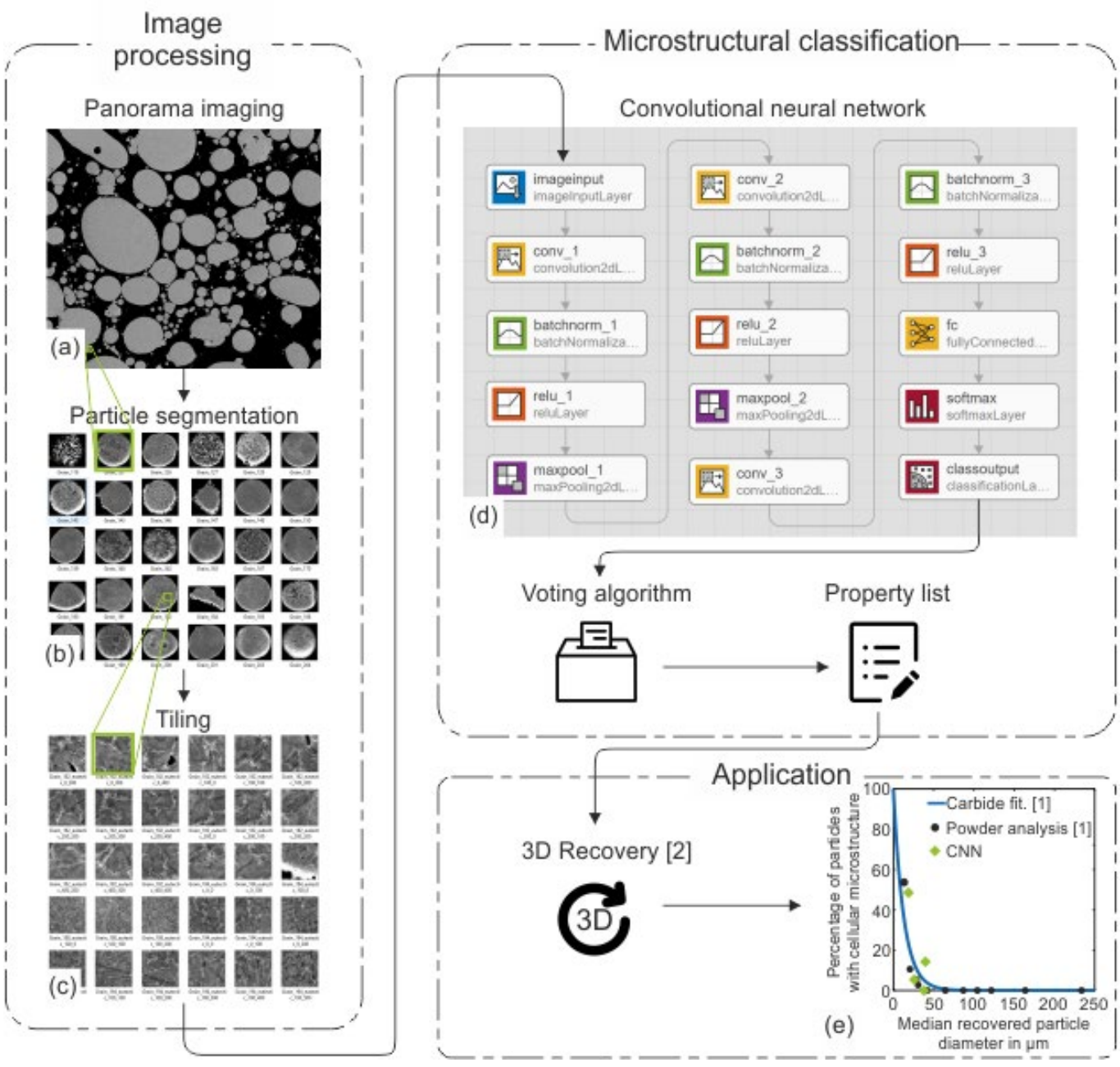


Figure 1: Proposed automatic characterization scheme.

At the Chair of Materials Technology (LWT), we dissect the relationships that permeate and define material selection and design. Correspondingly, the observation, description, and ultimate prediction of causal connections between processing and emerging macroscopic properties stand at the heart of our activities. This exercise is deductive and traditionally bases all associations on physical parameters and their interpretation. For instance, we routinely investigate the effects of manufacturing variables on the resulting mechanical properties, such as hardness, toughness, or fatigue life. To that end, the microstructure is the subject of intense examination, as it is ultimately responsible for the observed emergent behavior. Many of the scientific or technical questions that we strive to answer boil down to quantitatively studying the—sometimes subtle—effects of processing on the microstructure in terms of known or hypothesized thermodynamic and kinetic phenomena. Then, we scrutinize the microstructure as the thermomechanical interaction of its microconstituents to explain observed or expected measurable macroscopic properties.

Incorporating Data Science and Informatics into the rigorous framework of Materials Science and Engineering is the quintessence of Integrated Computational Materials Engineering and Materials Informatics. In Materials Engineering, the challenge lies in developing information-based principles that emphasize physical meaning. In other words, technological applications set feasibility boundary conditions that cannot be circumvented. Data-based methods require careful adaptation and implementation to ensure applicability and relevance in application-oriented research.

In this short article, we present one of our efforts toward increasing speed and reliability in repetitive microstructural analysis tasks through machine learning. Microstructural characterization represents a prime candidate activity because of three reasons: Firstly, micrographs are nothing other than data arranged in matrices. Secondly, large imaging campaigns produce heaps of images. And, thirdly, we can embed these automatic analyses directly in existing workflows without ever straying from standard physical descriptors.

Figure 1 presents the proposed approach. In a nutshell, we trained a convolutional neural network (CNN) to recognize the textures of the solidification structures in powder samples of the high-speed steel PM HS 3-3-4, depicted in Figure 2. Due to the extreme cooling speeds of the smallest steel droplets, they solidify into the very fine cellular microstructure exemplarily displayed in Figure 2(a). In contrast, Figure 2(b) shows a more massive droplet with regular eutectic solidification and the corresponding eutectic carbides. Telling these mechanisms apart is technologically relevant because they produce distinct carbide morphologies when the powder is processed in the hot isostatic press to a semi-finished product [1]. To complement the CNN, we added the particle segmentation and tiling steps and included a voting algorithm that democratically decides the solidification mechanism based on the decision on the individual tiles. These processing stages are detailed in Figure 1(a)-(c). With the trained CNN architecture in Figure 1(d), we are able to automatically ascribe each of the 1500 powder particles in one square millimeter area to either category in a matter of minutes. In Figure 2(e), we present a use case. We expanded each segmented particle's property list to include its equivalent area diameter besides the solidification structure. With the help of a three-dimensional size recovery algorithm [2], we could reproduce the trends published in [1] in a fraction of the original time.

At LWT, we actively explore original ways to generate new knowledge and efficiently improve existing exercises. As stated above, this brief account serves as an example.

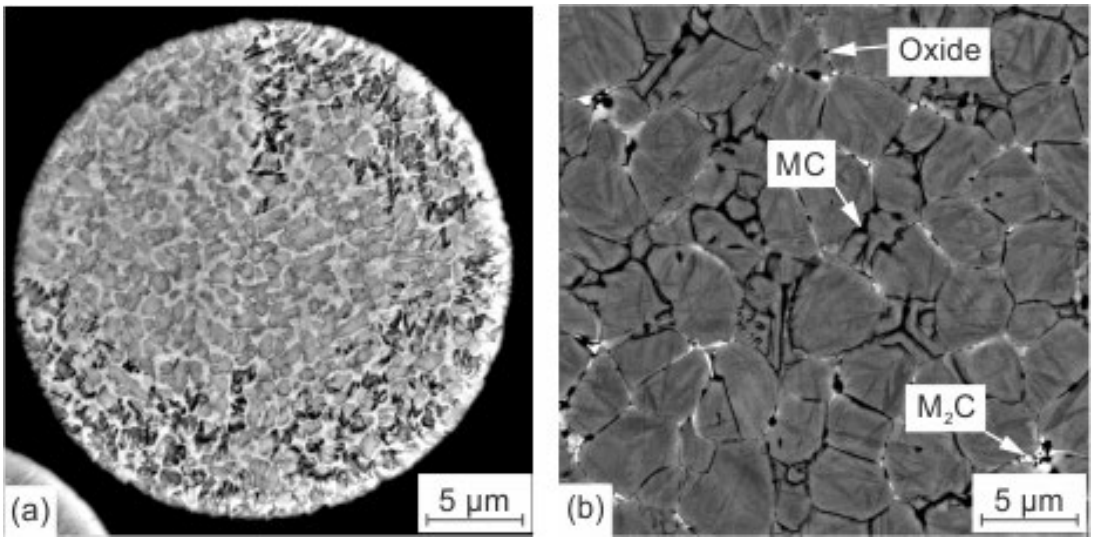


Figure 2: Characteristic microstructural solidification structures of the PM HS 3-3-4 powder droplets.

Santiago Benito, M.Sc.
Chair Materials Technology
santiago.benito@rub.de

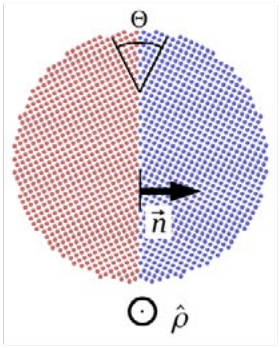
Prof. Dr. Sebastian Weber
Chair Materials Technology
weber@wtech.rub.de

REFERENCES:

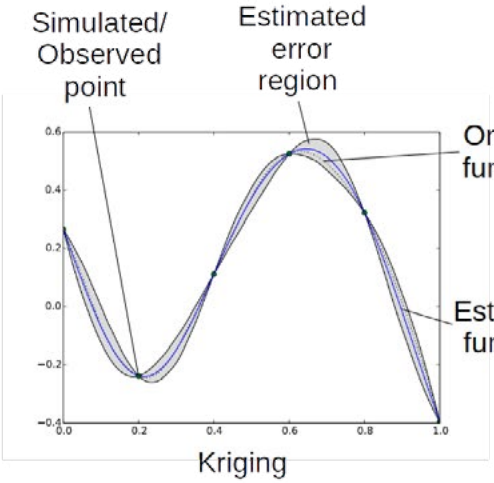
[1] S. Benito, et al., Materials & Design (2021), 208, 109925.
[2] S. Benito et al., Materials Characterization, (2019), 156, 109872.

WEB :

If you are interested to know more, visit our webpage or get in touch with the authors
↗ **www.wtech.rub.de**



Spherical sample



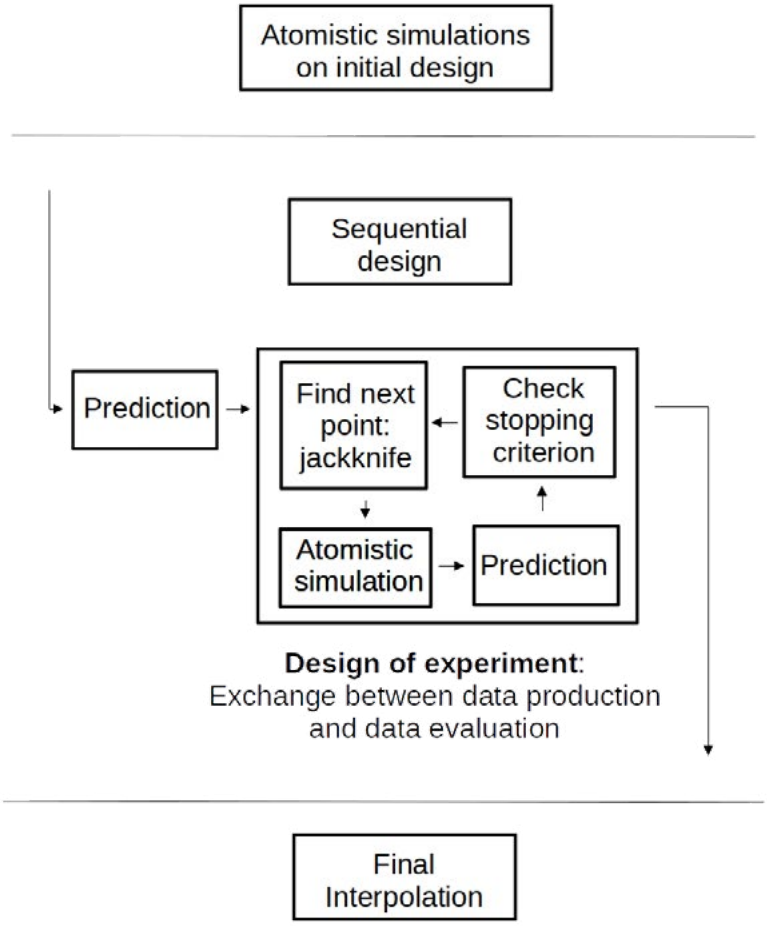


Figure 1: Overview over sequential design of experiment scheme.

ATOMISTICS MEETS STATISTICS

An Efficient Sampling of the Grain Boundary Parameter Space

The evolution of a metallic or multiphase microstructure during solidification or mechanical deformation largely depends on the structure and energy of the interfaces in the material. So do the resulting macroscopic properties, like the deformability, thermal stability, and strength. Consequently, substantial scientific effort is made to investigate grain boundary properties, and a valuable tool to do so are high-throughput atomistic simulations.

Nowadays high-throughput numerical simulations are standard in materials development, because they allow a systematic variation of material or process parameters. The most common approach to cover a broad range of these parameters in a short time is based on a regular, i.e. equidistant sampling of the parameter space, which keeps the automatisation of the workflow fairly simple. It has limited use however, when the properties of interest depend on several variables at the same time, i.e. a multidimensional parameter space has to be sampled, and the property of interest does not vary homogeneously in this space. This is the case for the energy of grain boundaries, which depends on the five macroscopic geometric degrees of freedom of the interface, defined by the rotation axis and angle and the inclination of the grain boundary plane, and exhibits deep local minima, so-called cusps.

For such cases a sequential design of experiment (DOE) can be beneficial, during which the next points in the parameter space to be sampled

are chosen based on the already available data [1,2] (see Figure 1).

The first important building block of the numerical recipe is the atomistic representation and optimisation of the grain boundary structure starting from the macroscopic degrees of freedom, and the calculation of the grain boundary energy. If the sampling scheme shall be valid for the complete 5D parameter space, including small angle and other general grain boundaries, the main challenge is to realise such a calculation without periodic boundary conditions. This can be achieved by using a spherical sample. The two half-spheres can be cut and oriented to represent any desired combination of misorientation axis, rotation angle, and plane inclination. By evaluating only the atoms in the interior of the sphere, surface effects are excluded.

After creating a first set of structure-energy data with atomistic simulations, a suitable interpolation method is required. The so-called Kriging (also referred to as Gaussian process regression in the presence of noisy observations) predicts the value of a function at a given point by computing a weighted average of the already determined values of the function in the neighborhood of the point. This approach is related to a regression analysis, but Kriging exactly interpolates through the existing observations. Furthermore, it provides a natural measure of uncertainty quantification for predictions at potential locations. In the sequential design, this information is used to create a list of candidate points for new sampling locations from the regions where the uncertainty is maximal. In these regions the points with the maximum expected variance compared to their neighbors will be chosen, in other words, points which are expected to be on a steep slope in the energy landscape.

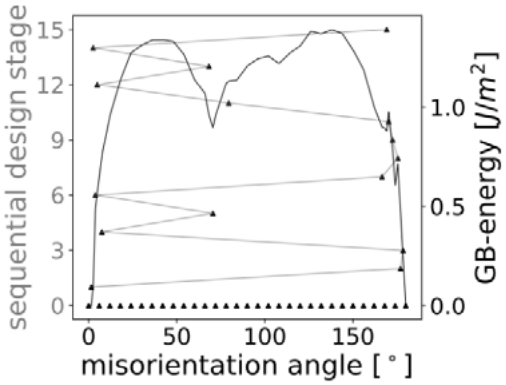


Figure 2: Sampling the energies a symmetrical tilt grain boundaries.

A typical sequence of experiments for a one-dimensional example – the energy of symmetrical tilt grain boundaries as a function of misorientation angle – is shown in Figure 2. It demonstrates the advantage of the sequential design, namely that cusps which are not included in the initial data are identified by the algorithm after a few steps only. This is particularly advantageous for more complex subsections of the 5D parameter space. As an example, Figure 3 shows the energy for different inclinations of the plane (given by the two angles that define the normal vector of the plane) of a small-angle grain boundary. The energy minima in such a subspace cannot be easily predicted from crystal symmetry, but are revealed by the DOE algorithm. The underlying strategy will be advantageous for any application with strong, localized fluctuations in the values of the unknown function.

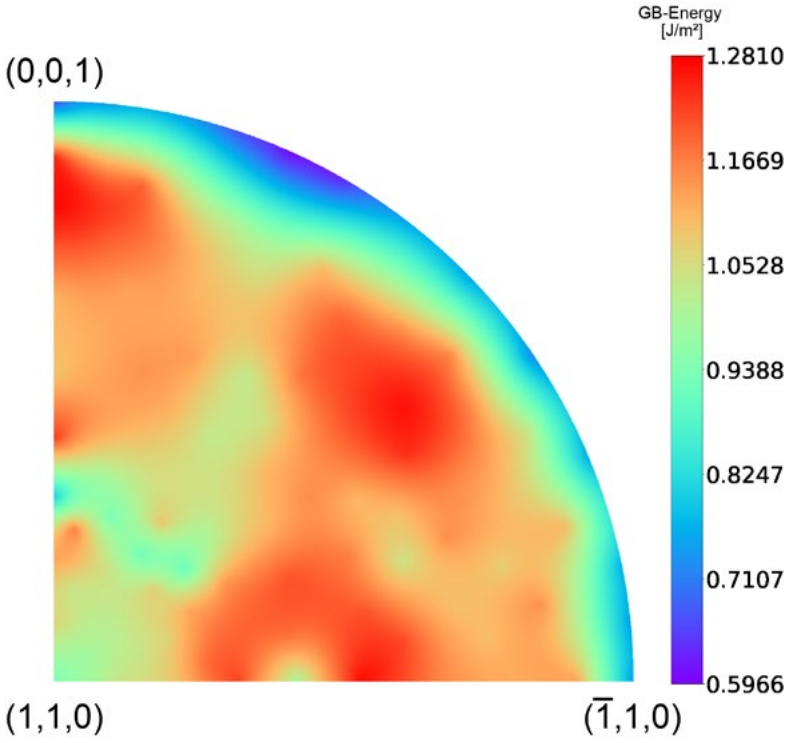


Figure 3: Fundamental zone of grain boundary energies of a low angle grain boundary with different boundary plane inclinations.

Dr. Martin Kroll
Chair of Stochastics and Machine Learning (Uni. Bayreuth)
Martin.Kroll@uni-bayreuth.de

Timo Schmalofski, M.Sc.
Mechanical Properties of Interfaces
timo.schmalofski@icams.rub.de

Prof. Dr. Holger Dette
Institute of Statistics
Holger.dette@rub.de

PD Dr. Rebecca Janisch
Mechanical Properties of Interfaces
rebecca.janisch@rub.de

REFERENCES:

[1] Dette, H., et al., Acta Materialia (2017), 125, 145-155.
[2] Kroll, M., et al., Advanced Theory & Simulation (2022), 5(7), 2100615.

THREE-DIMENSIONAL MICROSTRUCTURE RECONSTRUCTION FROM SURFACE EBSD MAPS

Reducing Effort by Micromechanical Simulations

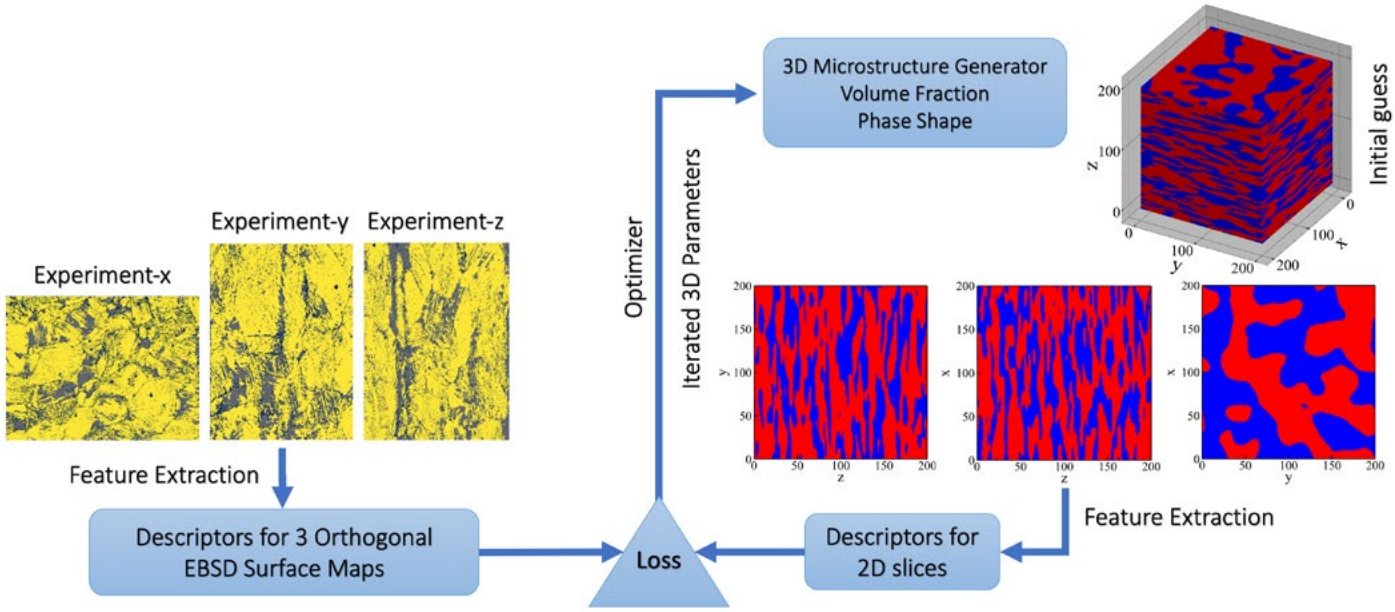


Figure 1. Representation of workflow to reconstruct 3D microstructures from surface maps.

The characterization of three-dimensional (3D) microstructures that captures the essential features of a given material is oftentimes desirable for determining critical mechanisms of deformation and failure and for conducting computational modeling to predict the material's behavior under thermo-mechanical loading conditions [1]. However, acquiring 3D microstructure representations is costly and time-consuming because standard microscopic procedures can only produce 2D surface maps. Hence, current state-of-the-art methods for 3D microstructure characterization are serial sectioning techniques or X-ray tomography. Both methods produce a truthful characterization of the 3D structure of an individual specimen but require a rather high effort both in sample preparation and in software-based reconstruction of the 3D structure. In this work, an alternative path is suggested to generate synthetic 3D microstructures that resemble real ones in a statistical sense with a severely reduced effort, as only 2D surface maps from three orthogonal surfaces are used. The method is based on an inverse procedure that generates synthetic 3D microstructures with arbitrary parameters and then compares these artificial surface maps with the real ones. In an iterative procedure, the parameters of the microstructure generator are optimized until the best possible agreement between the corresponding surface maps of synthetic and real microstructures is achieved. In this way, the statistical descriptors of the real microstructure are gained as they are represented by the converged input parameters for the microstructure generator. Advantages of this method are that, after the converged microstructure parameters are obtained,

different realizations of statistically equivalent 3D microstructures can be generated and also systematic parametric can be conducted by varying individual microstructure features. In this way, the influence of microstructure features on the material properties can be predicted by micromechanical simulations [2].

The work presented here focuses on microstructures of metastable austenitic steels where austenite and deformation-induced a-martensite co-exist at room temperature. The microstructure of this dual-phase steel is characterized by electron backscatter diffraction (EBSD) microscopy to produce three maps from orthogonal surfaces. The 3D microstructure reconstruction is performed in the way described above based on three EBSD maps from orthogonal surfaces of a dual-phase steel sample. As illustrated in Figure 1, the workflow for such reconstructions incorporates a microstructure generator tool from the open-source python package pyMKS to produce synthetic 3D dual-phase microstructures based on physical parameters, such as volume fraction and phase shape [3]. The input in Figure 1 is represented by three EBSD maps from orthogonal surfaces of a dual-phase steel (yellow: martensite, blue: austenite) from which a low-dimensional, yet representative vector of descriptors is extracted. The corresponding descriptors are generated from a synthetic 3D microstructure (red: austenite, blue: austenite) such that a scalar loss function can be evaluated to assess how similar the surface maps of real and synthetic microstructures are. In an iterative procedure, the parameters of the 3D microstructure generator are optimized such that the loss function becomes minimal. From the surfaces of the synthetic microstructure, 2D images are produced similarly to the surface maps of the real material. The primary challenge in minimizing the differences between real and synthetic surface maps lies in defining a proper loss function that quantifies the differences between surface maps in a physically sound yet numerically efficient way. In the present work, it is demonstrated that processing surface maps by spatial correlation functions, often referred to as 2-point statistics, and principal component analysis (PCA) results in a small set of unique descriptors that serve as a fingerprint of the 2D maps [4]. These descriptors encode the topological information of 2D maps in a compact format and can be used to characterize both experimental and synthetic surface maps. In this way, the differences between the two surface maps can be quantified and iteratively minimized, as seen in Figure 2. After convergence, the synthetic 3D microstructure accurately describes the experimental system in terms of physical parameters such as volume fraction (here: 15% austenite) and phase shapes (here: aspect ratio of 20:4:3 for martensite regions). Hence, the presented approach ensures that the 3D reconstructed sample and the associated 2D surface maps are statistically equivalent.

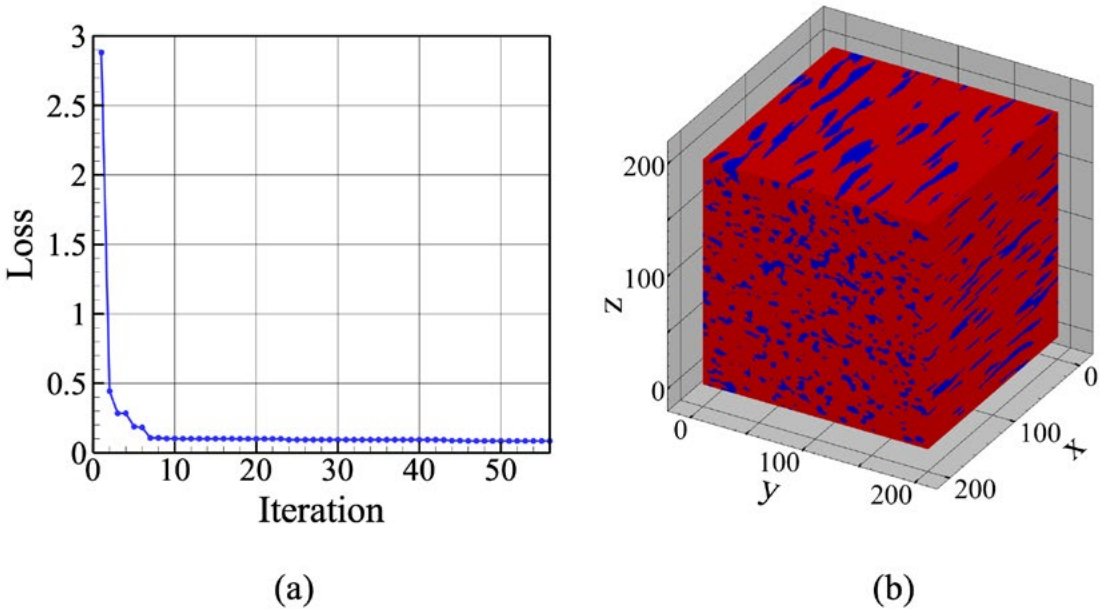


Figure 2. (a) Evolution of loss function during optimization procedure and (b) final synthetic 3D microstructure mimicking the experimental surface slides in an optimal way.

Golsa Tolooei Eshlaghi, M.Sc.
Department of Micromechanical and Macroscopic Modelling
Golsa.TolooeiEshlaghi@rub.de

Dr.-Ing. Gero Egels
Chair of Materials Technology
egels@wtech.rub.de

Prof. Dr. Sebastian Weber
Chair of Materials Technology
weber@wtech.rub.de

Prof. Dr. Alexander Hartmaier
Department of Micromechanical and Macroscopic Modelling
alexander.hartmaier@rub.de

REFERENCES:

[1] Benito, S., et al., Materials Characterization (2019), 156, 109872.
[2] Biswas, A., et al., Advanced Engineering Materials (2019), 21(7), 1900275.
[3] Brough, D. B., et al., Integrating materials and manufacturing innovation (2017), 6(1), 36-53.
[4] Niezgoda, S. R., et al., Integrating Materials and Manufacturing Innovation (2013), 2(1), 54-80.

ATOMIC CLUSTER EXPANSION FOR LARGE-SCALE ATOMISTIC SIMULATIONS WITH QUANTUM ACCURACY

Learning the interaction between atoms

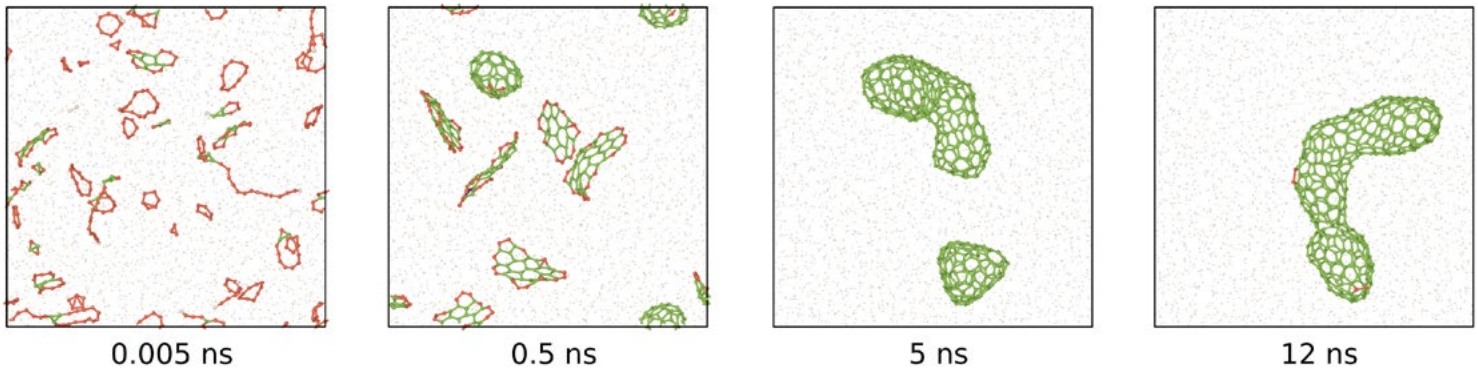


Figure 1. Nucleation and growth of fullerences in an argon atmosphere.

Electronic structure calculations have become a valuable tool for materials research and nowadays take a prominent place in many research projects. However, their high computational cost still presents a major limitation and simulations for only a few hundred atoms and picosecond time scales require access to supercomputers. This limits modeling of many important phenomena in materials science, chemistry, physics, and biology, for instance, predictions of phase diagrams and phase transformations, behavior of extended defects, diffusion, material degradation, etc. To be able to simulate these phenomena one has to invent effective interaction potentials that mimic closely the energies and forces from electronic structure calculations, but at a much lower computational cost. By adopting methods from machine learning (ML) and data science, tremendous progress could be achieved in recent years. To construct a ML interatomic potential, one starts by carrying out large numbers of automated high-throughput density functional calculations (DFT), typically for tens of thousands of different structures. The DFT database is then used to train the ML potential by varying its parameters to match the reference data as closely as possible. After careful validation, the ML interatomic potential is ready for large-scale simulations with millions of atoms and simulation times of nanoseconds.

We have recently developed the Atomic Cluster Expansion (ACE) [1] and demonstrated that it shifts previously established limits of ML interatomic potentials towards more accurate and numerically more efficient simulations [2]. The method is general and suitable for most elements and compounds across the periodic table. Efficient implementations for CPU and GPU hardware are available from ICAMS and have already been incorporated in prominent simulations codes, such as LAMMPS [2,3].

Figure 1 illustrates a long time-scale simulation of nucleation and growth of carbon fullerenes [4]. Initially, carbon atoms are distributed randomly in an argon atmosphere under high pressure and at high temperature. Forces between the atoms are predicted by ACE and the acceleration is integrated into a trajectory for each atom. From left to right, one first observes the nucleation of small carbon clusters that gradually coalesce and form fullerene molecules. At the end of the simulations only a single large fullerene molecule remains. The same ACE potential that predicts fullerene formation and growth can be used for diverse simulations of amorphous carbon, crack propagation in diamond, or defects in graphene, thereby providing a simulation tool that is able to cross the boundaries between materials science, chemistry, physics and biology.

Prof. Dr. Ralf Drautz
Department for Atomistic Modelling and Simulation
ralf.drautz@rub.de

REFERENCES:

[1] Drautz, R., Physical Review B (2019), 99, 014104.
[2] Lysogorskiy, Y., et a., npj Computational Materials (2021), 7, 97.
[3] Bochkarev, A., Physical Review Materials (2022), 6(1), 013804.
[4] Qamar, M., et a., https://doi.org/10.48550/arXiv.2210.09161.